(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 December 2000 (14.12.2000)

PCT

(10) International Publication Number
WO 00/75779 A2

(51) International Patent Classification[7]: G06F 9/46

(21) International Application Number: PCT/US00/15224

(22) International Filing Date: 2 June 2000 (02.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/137,568 4 June 1999 (04.06.1999) US

(71) Applicant *(for all designated States except US)*: IWITNESS, INC. [US/US]; Suite 2N, 2995 Wilderness Place, Boulder, CO 80301 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: LAMBERT, Francis, T. [US/US]; 1901 Spruce Street, Boulder, CO 80302 (US). ARCHER, Emory, Scott [US/US]; 4799 White Rock Circle, #E, Boulder, CO 80301 (US).

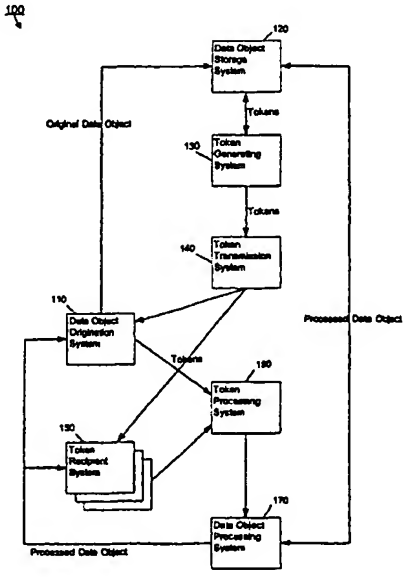
(74) Agents: SABETT, Randy, V. et al.; Cooley Godward LLP, One Freedom Square-Reston Town Center, 11951 Freedom Drive, Reston, VA 20190-5601 (US).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,

[Continued on next page]

(54) Title: TOKEN BASED DATA PROCESSING SYSTEMS AND METHODS



(57) Abstract: A token based data processing system is disclosed that uses data object tokens (DOTs) to manage the associated data objects within the data processing system. The numerous advantages of a token based data processing system stem from the ability to utilize a DOT separate from its associated data object. A data processing system according to the present invention can include a data object origination system, a data object storage system, a token generating system, a token transmission system, one or more token recipient systems, a token processing system, and a data object processing system. The token generating system can generate a DOT that contains metadata corresponding to the information that the token generating system receives regarding the data object, including its attributes, the parameters stored with it, and other information regarding its originating system and its recipient system. After being distributed by a token transmission system, token recipient systems can analyze and process the DOTs in lieu of the associated data objects.

WO 00/75779 A2

IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

TOKEN BASED DATA PROCESSING SYSTEMS AND METHODS

RELATED APPLICATION

The present application claims priority to Provisional Application No. 60/137,568, filed June 4, 1999, which is incorporated herein by reference in its entirety.

BACKGROUND

Field of the Invention

This invention relates to a data processing system containing diverse data objects and to methods for using those data objects. More specifically, the invention relates to a data processing system where the system utilizes data objects called "tokens" to perform processes related to, and upon, other data objects to which the tokens refer.

Description of the Related Art

The advent of electronic data interchange, the growth of electronic communications, the growth of the number of electronic communication users, the growth of the availability of information, and the growth in the size, number and types of data objects has greatly increased the requirements of both data storage and data communication bandwidth. Typical use of these data objects requires processing and transfer of the data object itself, often due to the complexity or sensitivity of the data content of the object, or to the limitations of the data systems used to manipulate the data object. These systems generally require transfer or transmission of the entire data object for a user to be able to edit, preserve, alter, or "own" the data on the person's

local computing system. Examples of these systems can include any that use e-mail attachments, ftp file transfer, and http file transfer.

Unfortunately, a modern data object storage event or a transmission event, such as an electronic document transfer via e-mail, may contain a large volume of data that one entity wishes to transfer to another entity. This type of transfer forces both the sending and receiving parties to expend substantial resources in terms of CPU usage, data storage, and transmission bandwidth at both ends of the transmission. The cost and lost productivity from these types of transmissions will increase as the number and size of these transmissions increase, as is the case in modern data communications.

Additionally, the originator of the transmitted data may wish to keep the transfer of the data secure and private. This requires sending the data using encryption, which requires additional CPU usage. The originator may also wish to guarantee that the transmitted data arrived exactly as sent, which requires that all of the data be sent, verified, and the verification returned to the originating system each time the data is transmitted. These security and authentication technologies must also be present on the receiving system. Prior to the data being released to the transmission network, the originator must have knowledge that the receiving system is capable of reliably authenticating, decrypting, and controlling access to the recipient's copy of the data object.

In many cases, the receiver of the data in this type of transmission is forced to receive all of the transmitted data whenever the originator sends it, regardless of whether or not it is convenient for the receiver at that time. An example of this is a mobile computer user being forced to download a large e-mail attachment over an expensive hotel data line so that they can retrieve all of their e-mail. Another example

is a wireless data transceiver being forced to download unnecessary data when only limited bandwidth is available.

In the case where a data object is transmitted to multiple recipients, the originator is required to transmit the data object multiple times, thereby expending valuable CPU and bandwidth resources in a the concentrated time period of transmission Also, each recipient is required to receive and store its copy of the transmitted object, thereby using bandwidth and data storage resources at each recipient. If the recipients are receiving diverse variations on a data object, the originator is required to create, store, and transmit the separate variations of the data object.

It is also possible that the volume of data in the transmitted data object is too large for the receiving system to efficiently store, or that numerous accumulated objects are too large for the system to efficiently store, and that it would be desirable to have another system with greater storage capacity hold the object and provide authorized access to the object only as needed.

It is also possible that the originator of the data would want to control the access and use of the object by the receiver, or create an audit trail of the receiver's use of the data object. The originator may wish that various receiving parties have access only to portions of the data object specific to the nature of the receiver. The originator or the receiver may desire to provide authentication relating to a data object without transmitting or storing the object itself.

Mechanisms have existed in the past for providing information about stored data objects. For example, a computer file (one type of data object) under the Disk Operating System (DOS) or the Windows operating system has certain flags associated

with it to indicate information about that file. The flags indicate, for example, whether the file is "read only", whether it has been archived, and whether it can be written. The UNIX operating system utilizes a similar approach of storing flags with files. However, the flags in all of these systems are rudimentary, and they are not easily extensible. The flags also tend to be stored as part of the object to which they refer and are therefore cumbersome to process without access to the file itself.

Other types of data objects contain elements that refer to other data objects within the context of the containing object. An example of this could be a filename combined with disk sector locations in a storage directory object, such as a file allocation table. Another example could be a hyperlink with addressing information contained within an HTML object. These object elements act as "pointers" to an object, but do not function outside of the context of the containing object. They also require a process to access the referenced object to perform processes or analysis on the referenced object. They do not act as a "surrogate" for the object, allowing processes to perform functions on an object without accessing the object itself.

Unfortunately, current data network transmission and storage systems fail to provide a robust, extensible, and universally compatible means for those users or systems transmitting and receiving data to efficiently use data objects in such as way as to conserve CPU usage, bandwidth usage, and storage through the use of smaller, more efficient objects containing metadata about the data object that can be processed separately from the data object itself. This smaller, more efficient object can contain metadata for use in access control, use auditing, redactability (or the filtering of content in an object), recipient and sender authentication, workflow logic, data integrity, and

timing of data delivery, among others. Consequently, a need exists for a data processing and transfer system that overcomes the foregoing drawbacks.

SUMMARY OF THE INVENTION

The present invention provides a data processing system that can allow a user to manipulate a data object through the use of one or more separate data objects comprised of metadata about the original data object. The metadata can represent many attributes of a data object, including, but not limited to, information about the content of the data object and the context of its occurrence. It can also represent information about searchable elements in the data object or information on the use history of the object. Some specific examples of metadata can include object identification, access parameters, authentication requirements and values (for both originator and user), encryption methods and keys, access history, integrity checking values, redaction values (or restricting access to certain elements or properties of the original data object), and token validity expiration period. In an embodiment, the DOT can also contain references (such as hyperlinks) to other locations where other metadata values can be found. Systems holding metadata values can also control access to and log use of those metadata values in other locations according to the contents of the token. A data processing system according to the present invention can combine diverse metadata into a discrete and processable data object called a Data Object Token (DOT) that refers to one or more data objects. The DOT includes metadata that provides information about a data object.

In one embodiment, a data object originator can submit a data object to a token based data processing system, which can optionally store the data object and generate

one or more tokens containing metadata of varying complexity referring to the stored data object. The one or more tokens can then be returned to the originator, or sent directly to token recipients, or stored separately in the data object storage system for later use. This Data Object Token, the DOT, can be used by the originator or by token recipients to interact with the data object that is optionally stored on a system storage device such as a central data depository attached to a network, or in another location specified in the DOT according to the protocol used to locate a data object, including a URL, a storage device identifier, or other data object locating protocols. The DOT user can submit the DOT to the token based data processing system to retrieve, view, alter, delete, or otherwise interact with the data object. The token based data processing system can analyze the metadata values contained in the DOT to determine use of the data object, including allowing access, processing authentication values, performing redaction, transmiting the object or portions thereof, destroying the data object, or otherwise manipulating the data object. It can also store and transmit metadata about data object as a smaller, discrete object that can be managed using native and pre-existent file and transmission processes.

One advantage of a token based data processing system is that it allows users of data objects to have access to and control over data objects without being required to expend local storage or bandwidth resources in the process. A DOT can be of very small size and still refer to a much larger data object in a remote location, while providing many capabilities to the local processes that would otherwise require the presence of the data object on the local system. A DOT can be very efficiently transmitted to multiple users, requiring much less storage and bandwidth than multiple transmissions of the referenced data object.

Another advantage is that the user of the DOT can retrieve, download, view, or otherwise process the referenced data object at their convenience by submitting the DOT to the token based data processing system. Another advantage is that the user can receive a DOT that allows them to decide upon the usefulness of, and securely retrieve, the referenced data object without requiring them to download the data object at time of originator transmission to them.

Furthermore, a token based data processing system can contain the authentication values and methods for both the originator and the DOT user that validate access and other rights to the data object. The token based data processing system can validate the DOT as it contains the originator authentication, as well as validate other parameters of the DOT based on the user authentication, before access, processing, or retrieval of the original data object is allowed.

The DOT can also contain data redaction information tailored to the nature of the DOT user. The DOT can provide the system information as to which portions of the data object the user will be allowed access. This relieves the data object storage system from the requirement of having to store and process redaction information for every potential user of the data object.

A further advantage is that the DOT can contain data object routing information for workflow and rules processing of data objects. The process can be more efficient and less bandwidth intensive if a token representing the workflow data object is passed around instead of the object itself.

Another advantage is that the DOT can append data object usage information to itself whenever it is used, and thereby become a self-contained data object usage audit

trail. This usage information could contain user identity and authentication, and time and type of use of use of both the token and the referenced object.

Yet another advantage is that the data object originator can stipulate in the DOT an expiration or starting time for processing of the referenced data object. This allows an originator to control the timing of various aspects of processing the stored data object, as well as avoid the need to provide a local copy of the data object to the DOT user.

Another advantage to using a DOT as a surrogate for a data object is that it avoids exposure of the data object to potential security and privacy problems inherent in the transmission of data over public networks such as the Internet. For example, even if a DOT is sent to the wrong recipient, or "sniffed" from the public network, the referenced object cannot be processed or retrieved since the unauthorized token possessor will not have the correct authentication to access the object.

Another advantage to using the DOT as a surrogate to the referenced data object is that the DOT can contain searchable elements of the original data object, such as keywords or text, and allow searching of the original data object, or a set of data objects, without requiring access to them. This allows searching to be distributed to the token user machine, thus conserving bandwidth and CPU usage at the object storage system.

Another advantage of using a DOT to retrieve an object is that an integrity value, such as a hash digest, of the retrieved data object can be compared to an integrity value in the token user's DOT to determine if the data object was received correctly.

Another advantage of using a DOT approach is that a token generating system can be embedded in an e-mail server, allowing the e-mail system to embed a DOT in

the place of large e-mail attachments. This conserves bandwidth and storage for both the sender and receiver.

Another advantage of DOTs is that they can be used as an electronic payment system due to their high level of security and ability to provide an audit trail. For example, the DOT can represent a portion of a pre-paid monetary amount represented by the referenced data object, such as a bank account. The payment DOT can be submitted to a token processing system that reduces the bank account balance by the amount embedded in the DOT, and then expires the DOT by placing its unique identifier on an expired DOT list. This would allow cashless payment in electronic business transactions, without exposing any sensitive financial information to a public network.

Another advantage of using DOTs is that they decouple the object storage technology from the object access technology, allowing object storage technologies to change without changing the way the stored objects are accessed by the users, and allowing the access objects, or DOTs, to remain valid and functional over long periods of time, even if the storage systems that they access change.

Another advantage of using DOTs is they it allow users to maintain representations of libraries of objects on a local system for off-line browsing, management, and processing.

Additional aspects and embodiments of the present invention will become apparent upon a review of the following detailed description and accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a functional block diagram of an embodiment of the present invention.

Fig. 2 depicts the process flow of a system according to the present invention.

Fig. 3 depicts the basic structure of a data object token.

Fig. 4 is a notational representation of the structure of one embodiment of a data object token.

Fig. 5 is a notational representation of additional elements in the structure of one embodiment of a data object token.

Fig. 6 is a command table for use in an embodiment of the present invention.

Fig. 7 is a flow chart of an embodiment of a token generating system.

Fig. 8 is a flow chart of an embodiment of a token processing system.

Fig. 9 is a flow chart of an embodiment of a token recipient system.

DETAILED DESCRIPTION

Fig. 1 depicts a block diagram of an embodiment of a token based data processing system 100 according to the present invention. Token based data processing system 100 can contain data object origination system 110, data object storage system 120, token generating system 130, token transmission system 140, token recipient system 150, token processing system 160, and data object processing system 170.

Data object origination system 110 can be any device or system that can be utilized to process information and produce data objects, including but not limited to such things as a personal computer, a mainframe computer, or an embedded computing system connected to a network. Data object origination system 110 can further include numerous processes that utilize data objects. For example, data object origination

system 110 could be utilized to create a data object and store that data object in data object storage system 120. Similarly, data object origination system 110 could also be utilized to download and store information, in the form of one or more data objects, from a data network. In particular, data object origination system 110 can utilize and store various data objects including, but not limited to, database files, electronic documents, e-commerce transactions (such as stock trades or merchandise purchases), electronic records (such as medical records or personal histories), multi-part records from document imaging systems, Web sites, sequences of HTTP data, or digitized video and audio data.

Following the creation or reception of one or more data objects in data object origination system 110, the data objects can be stored in data object storage system 120, which can be a local data storage system or a remote data storage system. For example, in the case of a personal computer, a local data object storage system 120 can include a hard drive, or a combination of a floppy disk drive and a floppy disk. Alternatively, a remote version of data object storage system 120 can include storage on a Local Area Network (LAN) or on a Wide Area Network (WAN) using such mechanisms as file servers or digital tape libraries.

In addition to receiving and storing the data objects, data object storage system 120 can expose the data objects to token generating system 130. Token generating system 130 can analyze and process the information about the data object, such as its attributes, the parameters stored with it, and other information regarding its originating system and its recipient system. Token generating system 130 can then generate metadata corresponding to that information, and create a data object called the Data Object Token (DOT) that contains the metadata associated with the data object. Token

generating system 130 can further process the DOT to embed data values that provide or ensure such things as security, privacy, reliability, or efficiency. Numerous means exist for providing these attributes, including encryption, digital signatures, hash digests, other authenticating technologies, compression, or other processing technologies. As one particular example, token generating system 130 can provide authentication, integrity, and non-repudiation information about the data objects using a public key digital signature. Upon successfully verifying the digital signature, the recipient of the digital signature can have the additional assurance of the authenticity of the originator, of the integrity of the data object, and that the data object creation cannot be repudiated.

Once processed, token generating system 130 can further present the DOT to data object storage system 120 and to token transmission system 140 for transmission to one or more token recipient systems 150. Data object storage system 120 or token transmission system 140 can store the DOT or a portion of the DOT for later use, thereby avoiding later reprocessing of the data object.

Token generating system 130 can read the object to be tokenized from data object storage system 120. Token generating system 130 can then determine from a token reference file the types and fields of metadata desired to be embedded in the token. A token reference file can be viewed as a template that contains rules governing the creation or transformation of tokens, including but not limited to encoding of the DOT (e.g. using XML, ASN.1, or comma-separated); required data and metadata fields; organizing structure of metadata fields; and mappings of data types and structures from the source data object to a token. Token generating system 130 can then read the metadata from the object and write it into a file that will become the DOT.

This metadata can include a unique identifier for the object, object location information, object hash digest values, or other metadata fields as required. Token generating system 130 can then determine from the token reference file if it is required to generate metadata from processing the referenced data object's content or metadata fields. Token generating system 130 can then read the content or metadata from the object, process it as required (such as, for example extracting keywords from content text, or reducing images for more efficient searching), and then embed the processed metadata into the token extension fields (see Fig. 3) in the DOT file. The metadata can be written in popular data formats, such as XML or comma delimited fields.

Token generating system 130 can then determine from the token reference file whether token generating system 130 is required to retrieve metadata external to the object, such as the date, a stock price, a digital certificate, encryption keys, other authentication data, or other data as is available to token generating system 130 and required by the token reference file. Token generating system 130 can then embed the externally retrieved data into the appropriate fields in the DOT file. It can also use predetermined time stamp and authentication fields, as well as token extension fields as required. Token generating system 130 can also embed additional metadata into the token file extension fields, such as token validity period, redaction values, or workflow logic to be used when referencing the original data object. Token generating system 130 then can embed its version number in the token file to allow later processing by token processing system 160 and, particularly, to differentiate among different versions of tokens.

Token generating system 130 can then digitally sign the DOT file by methods well known in the art of digital signing, including without limitation, encrypting, with a

private key from a public key infrastructure (PKI) key pair, a combination of a hash digest of the token with a digital certificate, and embedding the digital signature data into the token file. Token generating system 130 can then encrypt the entire token file using well known encryption algorithms, such as triple DES or the RC5 encryption methods.

The DOT file generated by token generating system 130 can then be stored in the data object storage system 120, in the data object generation system 110, or in token generating system 130 itself to await further processing.

Once token generating system 130 creates a DOT, token transmission system 140 can read the DOT from token generating system 130 and determine from a token transmission reference file whether or not the DOT is required to be transmitted to a token recipient system 150, or stored for later use. A token transmission reference file, similar to a token reference file, can contain template rules, such as recipient types, required transmission metadata, and required authentication, that would govern the transmission of DOTs. If the DOT is to be transmitted, the token transmission system 140 can embed information on accessors to the data object in the DOT as required. This information can be used by token processing system 160 to determine the authorization and access rights of the submitted DOT. The token transmission system 140 can embed multiple object accessors values to allow use of the DOT by multiple users or systems. The DOT can then be encrypted in the public key of the transmission destination system, and can then be securely and reliably transmitted to one or more token recipient systems 150 using either a standard communications protocol, such as FTP, e-mail attachment, or HTTP, or by using a proprietary communications protocol.

A token recipient system 150 system can include a system capable of receiving, storing, or processing a DOT. The token transmission system 140 can determine addressing and routing information for the destination token recipient systems 150 from a token transmission reference file, or from the token reference file contained in token generating system 130. In one embodiment, data object origination system 110 can also be a token recipient system. In an embodiment, token transmission system 140 can further verify and log reception of the DOT by one or more token recipient systems 150.

Token recipient system 150 can receive, store, and optionally process or otherwise condition a DOT. Processing on a DOT could include searching on metadata or content from the object embedded in the DOT. This searchable metadata or content could include, for example, text, keywords, or images from the original data object. Other types of processing that could be performed on a DOT by the token recipient system include embedding recipient reception information, such as time stamp and digital signature, appending additional accessors to the referenced object if the recipient system is so authorized, or forwarding the DOT to another token recipient system. The optional processing performed by token recipient system 150 can include other types of analysis and processing of the information contained in the DOT, including its metadata. Since token recipient system 150 can perform its analysis and processing on the metadata contained within the DOT (which is acting as a surrogate for the data object itself), no resources within data object storage system 120 would need to be used.

For example, in one embodiment, token recipient system 150 can contain an e-mail computer application that receives DOTs associated with e-mail messages. In this

case, each message could be a data object. The metadata contained within a particular DOT that is associated with a specific e-mail message can contain information about the e-mail message, including but not limited to such things as the date, time, subject, sender, and, in particular, any attachments associated with the e-mail message. The DOTs can be used by a token recipient system to search for a particular e-mail message or for a group of messages. However, this search could be done solely within the token recipient system using metadata embedded in the DOTs associated with the data objects (i.e. the e-mail messages), not on the messages themselves, which can be stored on a remote data object storage system 120. Also, using this approach, any attachments to the e-mail message would only need to be stored in the data object storage system, not sent to a token recipient system nor stored at the data object origination system.

Furthermore, in one particular embodiment related to e-mail, a token generating system can be embedded in an e-mail server to allow automated storage and tokenization of e-mail, including, without limitation, any of its components such as header, body, or attachments. Token generating system 130 embedded in the e-mail server can read the e-mail objects controlled by the server, generate a DOT for the e-mail object, replace the original e-mail object with the DOT, and store the e-mail object in a more efficient data object storage system location, such as a digital tape library. The e-mail user can then use the DOT from within the e-mail client application to access or process the e-mail data object. Searching for e-mail content within a group of e-mails can be accomplished by searching the search metadata embedded in the DOTs in the e-mail client application, thereby avoiding the CPU and bandwidth costs associated with searching the group of e-mails on the e-mail server or at the data object storage system. This process of substituting DOTs for data objects such as e-mails can

be applied to many types of data objects, including without limitation, scanned or imaged documents, electronic documents, database tables, electronic spreadsheets, web sites, and other types of data objects specified by the user of a token based data object management system.

Upon the completion of a search and a determination that one or more specific e-mail messages, or other type of data objects, need retrieval or processing, token recipient system 150 can submit one or more DOTs to token processing system 160 which can then cause data object storage system 120 to be accessed and the referenced data objects to be processed according to the instructions in the DOT, or transmitted to a data object processing system 170 as requested in the submitted DOT.

After token recipient system 150 submits a DOT to token processing system 160 to retrieve a data object from data object storage system 120, token processing system 160 can interpret the metadata contained in the DOT. Using the metadata, token processing system 160 can submit processing commands to data object processing system 170 according to the metadata contained in the DOT. These commands could include, for example, transmit, redact, alter, or destroy a data object as shown in Fig. 6.

Alternatively, a token recipient system can submit a DOT to a token processing system to retrieve a set of DOTs associated with the submitted DOT. This referenced set of DOTs can reference data objects or other DOTs, or a combination thereof, as required by the system.

Data object processing system 170 can interact with the data object stored in data object storage system 120 according to the commands submitted by token processing system 160. For example and without limitation, in response to a particular

command, data object processing system 170 can transmit, redact, alter, or destroy the data object stored on the data object storage system, according to the authorizations embedded in the submitted DOT, and as interpreted by the token processing system 160, which can pass the commands to the data object storage system 120.

In another embodiment where data object storage system 120 contains data object processing system 170, data object storage system 120 can process the referenced data object according to the instructions embedded in the DOT. For example, if data object storage system 120 is instructed to transmit the object, it can place the object into a data transmission process as is commonly known to the art, or it can encrypt the object using an encryption value submitted with the token, such as the destination system public key, or it can transmit the encrypted object to a system different from the system that submitted the token, or it can transmit the object to multiple systems. Similarly, data object storage system 120 can redact the stored object by recomposing the object prior to transmission, according to redaction rules embedded in the token. An example without limitation of this can be the removal of personal information from a data object for transmission of the data object from within the European Community to a destination outside of the European Community, thereby complying with the European Privacy Act, or other regulations that require redaction of information in transmission or storage.

Data object storage system 120 can additionally alter the data object by changing its retention information, its authorized accessor list, its storage format or location, or even its content if required by the system. Data object storage system 120 can even destroy the data object according to the processing instructions embedded in the submitted token.

Fig. 2 depicts a process flow diagram of an embodiment of the present invention. In data object origination system 210, a data object can be created. Once created, that data object can be transmitted to data object storage system 220. Once received, data object storage system 220 can process the data object so as to enhance the security and reliability of the data object storage process by, for example, performing encryption, compression, hashing, or mirroring, or by digitally signing the data object. For example, the well known DES encryption algorithm can be used to encrypt the data object. In addition, the well known RSA public key digital signature algorithm can be used to digitally sign the data object in data object storage system 220. Following the processing performed on the data object by data object storage system 220, the data object can be stored. The data object storage system can then place the processed object where it will be read by token generating system 230.

Token generating system 230 can generate tokens in standard data formats. For example, token generating system 230 can generate XML-based tokens based on the content of the data objects. To do so, token generating system 230 first reads the stored data object in data object storage system 220. It then analyzes the data object and the associated token reference files. The token reference files can contain instructions and information that can allow the token generating system to determine how the DOT will be constructed from the data object. Once analyzed, token generating system 230 can generate the necessary XML tags to be placed in the DOT for the particular system. In other embodiments, a token generating system can also, for example, generate fixed length fields, comma delimited fields, and other data formats that are commonly used by computer applications. In general, a token generating system can produce any type of DOT that may be needed for a particular data processing system.

In addition, token generating system 230 can encrypt, compress, hash and otherwise process one or more DOTs so as to enhance the security and reliability of the DOTs. Similar to the processing that can be done on the data object itself, the well known DES encryption algorithm can be used to encrypt the DOT. In addition, the well known RSA public key digital signature algorithm can be used to digitally sign the DOT in token generating system 230.

Furthermore, in an embodiment of the present invention, token generating system 230 can embed in the DOT time relevant attributes of the DOT, the data object, or the token recipient system's 250 ability to manipulate it. These time relevant attributes could include, for example, a time stamp of the DOT generation time, expiration of the DOT, the data object, or authentication metadata as determined by the contents of the token reference file. These time relevant attributes could also include "no access until" information that could not allow the token submitting entity to manipulate the data object until a certain time or event.

In an embodiment, once token generating system 230 has generated a DOT, token transmission system 240 can sense the DOT and can cause token generating system 230 to pass the DOT to token transmission system 240. To do so, token transmission system 240 can poll the system holding the DOT and retrieve the DOT in a data "pull" operation once it is available. Token transmission system 240 can then determine the destination for the DOT, followed by determining the methods to be used for the delivery of the DOT, according to mechanisms such as the transmission reference file in the token transmission system, or from the token reference file in the token generating system, or from a workflow logic file. Token transmission system 240 can then transmit DOTs using any well known data transmission technologies,

including, without limitation, the FTP protocol, the HTTP protocol, the SSL protocol, a virtual private network (VPN), or Secure Shell tunneling in the Unix operating system. In another embodiment, a token transmission system can send the DOT as an attachment to an e-mail message. Alternatively, a token transmission system can send a notification to a token recipient system of a token pending transmission to which token recipient system 250 must respond to release or access the token, and a token transmission system 240 can then issue a receipt describing the token delivery event.

After receipt of a DOT from a token transmission system, another token recipient system, or any other means of transmitting a token, such as in an e-mail as an attachment, token recipient system 250 can authenticate the DOT using any authentication techniques that may have been applied by token generating system 230. Once authenticated, token recipient system 250 can store the DOT and subsequently submit the DOT to token processing system 260 for data object manipulation.

In addition to submitting the DOT for data object manipulation, a token recipient system can also add new metadata to the received DOT, or alter the DOT according to the token recipient system's authority to do so as determined by the authorization metadata embedded in the DOT, and optionally transmit that updated DOT to another token recipient system or to the token processing system 260. This can allow metadata within a DOT to accumulate without having to access the associated data object from storage. A data processing system utilizing this type of DOT approach could undergo a change of the native platform and data formats of data object storage system 220 without impacting the capabilities of pre-existing DOTs used to manage and process the data objects within the system, provided the new native platform and data formats remained compatible with the required versions of the DOT structure.

In one embodiment, workflow logic can be embedded within the token or can be captured in an external file. Workflow logic refers to workflow processing rules stored and interpreted both inside and outside of a DOT, and that can affect both the internal operations of the token based data processing system and the overall workflow distribution patterns of tokens and data objects within an organization.

For example, workflow logic could be used by an organization to manage the workflow associated with a particular project. Workflow routing instructions could be embedded in the DOT if it is advantageous for the logic to be contained in the DOT, or the DOT could reference an external workflow logic file that could be dynamically updated. The DOT can reference an object or a set of objects intended to be the subject of a workflow process. This can allow the workflow system to distribute metadata about workflow-related data objects without expending the CPU and bandwidth to distribute the actual data object(s) themselves. The token recipient system 250 itself can execute the workflow logic embedded in the DOT, such as forwarding a DOT automatically to another token recipient system if the referenced data object is not accessed within a certain period of time, or other similar workflow logic.

Upon receiving the DOT from token recipient system 250, token processing system 260 can perform on the DOT any type of processing that might be necessary. For example, the DOT can be decrypted, uncompressed, or otherwise processed to restore it to its original form in response to security and reliability processing performed by the token generating system. The type of processing required can be determined by token processing system 260 by reading the metadata embedded in the DOT, such as the version number, or other metadata indicating how the DOT has been previously processed.

As another example, token processing system 260 can authenticate the entity submitting the DOT using digital certificates, recipient supplied passwords, biometric measurements, or any combination thereof, according to values contained in the token. Alternatively, the entity originating the DOT can be authenticated using digital certificates, recipient supplied passwords, biometric measurements, or any combination thereof, according to values contained in the token.

In yet another embodiment, token processing system 260 can create and append metadata to the DOT that can describe the use of the DOT and the data object. This metadata can be used create a self-contained audit trail within the token. For example, this self-contained audit trail can be comprised of use metadata and authentication values embedded into the token during a workflow execution by the associated token recipient systems.

Token processing system 260 can, in another embodiment, process metadata that describes the token submitting entity's access rights to the referenced data object. Those access rights can be used by data object processing system 270 to control access to all or portions of the data object. Metadata can also be used to describe the time attributes of a DOT, the referenced data object or the authenticating metadata. Token processing system 260 can interpret this time-based metadata to determine proper usage of the token and the associated data object.

The results of the token processing done by token processing system 260 can be passed as commands to data object processing system 270. Upon receiving one or more of these commands, data object processing system 270 can authenticate the command using any well known authentication technique already described herein. Upon successful authentication, data object processing system 270 can determine what

action needs to be taken to respond to the command. For example, destroying or deleting the referenced object from the object store, or appending the audit trail information embedded in the submitted DOT are actions that can be taken in response to a command.

One frequent action by data object processing system 270 can be the accessing of one or more data objects in data object storage system 220. If data object processing system 270 determines that the proper conditions have been met, including, for example, appropriate authentication information, one or more data objects can then be accessed in data object storage system 220. Once accessed, the requested actions can be performed on the accessed data objects by data object processing system 270 and those actions can also be logged by data object processing system 270. Finally, any results of the processing performed by data object processing system 270 can be passed back to token processing system 260. In another embodiment, data object processing system 270 can transmit one or more data objects to the token submitting entity. Prior to doing so, data object processing system 270 can optionally encrypt, compress, hash, sign, or otherwise process the data object to enhance security and reliability in transmission and delivery.

Figure 3 depicts a DOT 300 in accordance with the present invention. The basic DOT structure consists of a version field 304, token information 308, object accessors 328 , token and referenced object use information 352, and optional token extensions 368.

Version field 304 indicates the version of the data object token protocol to which DOT 300 conforms. The inclusion of version field 304 permits a token processing system to properly parse the DOT.

Token information 308 contains the fields necessary for a token processing system to assess the validity of the DOT and the data object to which the DOT is associated. Data object identifier 312 can contain a unique value or character string that identifies the data object to which it refers. Validity information 316 allows a system to make the determination of whether the DOT is valid. For example, validity information 316 can contain date and time information, such that a validity period can be established. Alternatively, validity information 316 can contain a hash digest of the object and the DOT that can be used for purposes of determining data integrity of a received object or DOT.

Data object and DOT authentication information 320 can contain any type of data that can be used to authenticate the data object or the DOT itself. In an embodiment, data object authentication information can contain a public key from a PKI key pair. In another embodiment, data object authentication information 320 can contain a digital signature. The authentication information can be used to authenticate the system creating the referenced data object, the system generating the DOT, or the system submitting the DOT to a token processing system or token recipient system. Timestamp 324 can provide information to an entity about the time of creation of the token. A data object originator could use the timestamp, for example, to prove that a data object representing a transaction was created at a specific time on a specific date.

The next main component of the DOT is object accessor 328. This can provide information about the entity or entities permitted to access and use the data object referenced by the DOT. In an embodiment, there may be multiple instantiations of object accessor 328 for each permitted accessor of the data object. Each object accessor 328 can contain information about the accessor 332, access rights 336,

authentication requirements 340, accessor authentication information 344, and timestamp 348. In an embodiment, the object accessor metadata can contain a list of workflow destinations upon which workflow distribution processing can be based.

Accessor 332 can contain information that identifies the entity being given rights to access the data object. This entity identifier could identify a system, a process, a class of processes, a specific user, a group of users, a reference to a dynamically alterable accessor list, or the general public. For example, accessor 332 can contain a name, an e-mail address, and a unique identifier. In another embodiment, accessor 332 can contain a digital certificate.

Access rights 336 can provide information to the token processing system regarding the rights that the accessor identified by accessor 332 has to the data object to which the DOT refers. Access rights 336 can include, for example, general rights, such as those used to access, download, modify, or destroy a data object. In addition, access rights 336 can include other, more specific rights that are unique to the particular data processing system. For example, a DOT used in a stock trading system might contain rights particular to a stock trader, such as buy unrestricted, buy limited, sell unrestricted, or sell limited. In another example, access rights can also stipulate authorization of the token recipient system to alter the metadata in the token, or the authorization to append audit trail metadata, as required by the system. Authentication requirements 340 provide information about the type of authentication required to be used by the token user to prove that it is the entity identified by accessor 332. Numerous, well understood, technologies exist to provide authentication in several different formats. For example, a particular data processing system may require something to be known by the token user, such as a password or personal identification

number (PIN). Alternatively, a data processing system may require the user to have an object, such as a smart card or a hardware "dongle". Finally, a data processing system might require physical information from the token user (i.e. a biometric), such as a fingerprint, a retinal scan, or a voice print.

Accessor authentication information 344 provides information about the entity that created the object accessor entry. This can be used to audit the granting of rights to access the referenced data object.

Timestamp 348 can provide information to an entity about the time of creation of the object accessor 328 metadata. A token processing system could use the timestamp, for example, to prove that the authentication information 344 was historically valid at the time of its creation. For example, if a digital certificate is used as authentication information 344, time stamp 348 can be used at a later time to determine if the digital certificate had expired or been revoked at the time of object accessor 328 metadata creation.

The next component of DOT 300 in an embodiment of the present invention, token use information 352, contains information about the use of the DOT and use of the data object to which the DOT refers. Token use information 352 can include information on the type of access requested 356 by the token user. Once authenticated using access authentication information 360, a token processing system can utilize token use information 352 to determine how it should process the object referenced by the DOT. This could include uploading the object to the token submitter, changing data in the object, or destroying the object.

One or more token extensions 368 allow the DOT to provide additional information beyond the standard DOT including, but not limited to, such things as

redaction instructions and audit trail data related to the associated data object. The format of each token extension 368 can consist of extension type 372, extension data 376, and extension authentication information 380. Within token extension 368, extensible metadata can be stored that can then be interpreted by the token recipient system or the token processing system. Some examples of the use of token extension 368 include DOT expiration values for a particular system, searchable content metadata values, workflow logic, additional (system specific) access rights, externally retrieved metadata, and any other fields that may be added by a token recipient system.

Fig. 4 and Fig. 5 provide a description of one embodiment of a DOT that can be associated with a Verified Data Object (VDO), as implemented by iWitness, Inc. of Boulder, Colorado in its electronic records management products. The descriptions in Fig. 4 and Fig. 5 use the well known Abstract Syntax Notation 1 (ASN.1) that is defined in "CCITT Recommendation X.208: Specification of Abstract Syntax Notation One (ASN.1)", published in 1988. The ASN.1 descriptions in Fig. 4 and Fig. 5 depict a Verified Data Object Token (VDOT).

Fig. 4 depicts the basic VDOT structure 410, which includes the two main components of tokenInfo and accessors. The tokenInfo component can provide information specific to the VDO, while the accessors component can provide information about the entities that may access the VDO. As shown in 420, the tokenInfo component can include globally unique object identifier vdo_id, the interval of time during which the token is valid, a digital signature and certificate identifying the creation event, and a timestamp. The vdo_id element can be used to uniquely identify the particular VDO to which the VDOT refers and can contain a randomly generated ASCII character string that had been verified as being unique. In a different

embodiment, the vdo_id could be a consecutive number applied to a particular VDO amongst a large set of other VDOs. The validFrom and validTo components together provide the validity period for the VDOT user to access the particular VDO. In an embodiment, if the UniversalTime value within the data processing system utilizing the VDO is less than the validFrom time or greater than the validTo time, the entity submitting the token in an attempt to access the associated VDO would be denied that access.

The signerInfo component provides the digital signature and information about the signer who applied the digital signature used to verify the VDO. In an embodiment, signerInfo can be the signer information defined in "PKCS #7: Cryptographic Message Syntax Standard: An RSA Laboratories Technical Note", Version 1.5, Revised November 1, 1993, which specifies, amongst other things, the signer's digital certificate, the algorithm used to sign the VDO, and the digital signature itself.

The accessors component within VDOT basic structure 410 specifies a collection of entities along with their access rights and authentication requirements for the specified VDO. As shown in accessor information 430, those entities can be identified in a table consisting of a set of DataObjectTokenAccessor elements, wherein each DataObjectTokenAccessor element can contain a PKCS #6 certificate, the access rights needed to access the VDO, the authentication requirements for the particular VDO, the PKCS #7 signer information, the PKCS #6 certificate of the signer of the VDOT, and a timestamp.

Access rights listing 440 for a particular VDOT can provide information on how a particular entity may use the VDOT in an embodiment of the invention. Only entities identified by the distinguished name and serial number from their enclosed PKCS #6

certificate can make use of the VDOT, including, for example, the ability to retrieve the corresponding VDO. Those entities are limited to their corresponding access rights, which are signed by the entity responsible for their assignment.

Authentication requirements 450 for a particular VDOT provides information on how a particular entity must authenticate itself prior to its use of the VDOT. Any use of the token, whether for VDO retrieval or communication to another entity, may require the using entity to satisfy one or more authentication requirements, such as knowledge of a password, a physical match of biometric data, or an external key. Additionally, each entry in the accessor table can be signed by its creator.

Fig. 5 depicts an ASN.1 description of the token use information 510 element and the token extension 520 element in accordance with an embodiment of the present invention. Token use information 510 can contain request information that indicates how a VDOT user would like to use the corresponding VDO and authentication information that allows the token processing system to determine the validity of the request. Token use information 510 can be included upon initial generation of the VDOT or can be appended at some later time after VDOT creation.

In an embodiment of the invention, one or more instantiations of the optional token extension 520 element can be included in a VDOT. In a VDOT, extType can specify the extension type of this particular extension.

Fig. 6 shows a command table implemented in one embodiment of the present invention that contains commands for performing various operations on a data object. A data object processing system can use the information shown in 610 to execute the commands on the specified data object. Transmit object command 620 can result in a data object being transmitted to the destination specified by the destination argument,

and can include information on authorization, transmission time, recipient requirements, and transport mechanism. Destroy object command 630 can result in the destruction of a data object, and can include information on authorization, destruction time, and disposition rules. Copy object command 640 can result in the copying of a data object from one location to the location specified by the destination argument. Alter object command 650 can result in one or more modifications to the data object, and can include information on the module to perform the operation, a reference to the data object, the actual alteration to be performed on the data object, and any necessary parameters. Redact object command 660 can result in certain information within a data object being removed while the rest of the data remains, and can include the module to perform the redaction, a reference to the data object, the redaction to be performed on the data object, and any necessary parameters. Finally, object encryption command 670 can result in the encryption or digital signing of a data object, and can include the operation to be performed on the data object, the method or algorithm to be used, the key to be used to perform the cryptographic operation, and whether to just perform a hash on the data object for data integrity purposes.

Figure 7 depicts a flow chart of the processing that can occur in a token generating system utilized for workflow processing. In step 704, the token generating system can receive a DOT generation request, which contains request attributes specifying the data object to be analyzed, the token reference file that defines the structure and required contents of the generated DOT, and possibly other control data. In order to ensure the security of the token generating system, and thus the security of the DOT and its future use in other systems, the DOT generation request is validated in step 712 through one or more of the various authentication means available including,

but not limited to, public key digital signatures. Once the DOT generation request is validated, the data object to which it refers is analyzed in step 716 to determine the viability of the request. In particular, as shown in the substeps in 720, the DOT generation request is checked against workflow logic (if any), the data object is then checked against workflow logic carried in the DOT generation request (if any), and finally, the data object is assessed to determine if it contains the necessary data/metadata to fulfill the request according to the token reference file. If the request cannot be fulfilled, a result indicating this can be returned in step 736.

If the authentication and subsequent analysis succeed, the token generation system can proceed to generate the requested DOT in step 724. As shown in substeps 728, all requested metadata that is derived directly from the data object can first be retrieved and attached. Then, all requested metadata that is derived external to the data object (if any) can be attached, possibly including workflow logic defined in the token request file. Finally, any requested DOT access rights can be constructed and attached. The newly formed DOT can then be digitally signed in step 732 by the token generation system, and possibly encrypted and enveloped following the Cryptographic Message Syntax as defined in PKCS #7. As shown in step 736, the DOT can then be returned synchronously to the originator of the DOT request and/or asynchronously transmitted through a token transmission system (as shown in the substeps in 740). Control could then be returned to the token transmission system in step 744.

Figure 8 depicts a flow chart of the processing that can occur in a token processing system utilized for workflow processing. In step 804, the token recipient system can send a processing request. In step 808, the token processing system can receive a processing request, which can contain, as shown in 812, a request DOT and

request attributes specifying one or more operations to be performed on the data object that the DOT references. Before processing, the token processing system can validate the public-key digital signatures on the DOT in step 816, to ensure the authenticity of the DOT and the authenticity of the requestor. Once validated, the token recipient system can open the cryptographic envelope of the DOT in step 816 (if it is encrypted) with the private-key that corresponds to the public-key used in securing the DOT for transmission. This ensures that the token can only be used by the intended token processing system.

Once validated and decrypted, the token processing system can check the token-embedded or external file workflow logic in step 820 and in the substeps shown in 824. The token processing system can then interpret the processing request according to the workflow logic of the token in step 828 in order to cause the data object processing system in 836 to perform the requested operations in step 832, such as fetch data/metadata, update metadata, destroy object, and transmit object/token, all of which can be performed on the data object within the data object processing system.

Results of the data object operation can be validated by the token processing system in step 840 to ensure their authenticity and conformance to the requirements of the request. The results can then be returned in step 844, and in particular, as shown in substeps 848, the results can be returned synchronously to the originator of the processing request and/or asynchronously transmitted as an updated token through a token transmission system. Once the results have been returned, control can return to the token transmission system in step 852.

Figure 9 depicts a flow chart of the processing that can occur in a token recipient system utilized for workflow processing. The token transmission system can

send a DOT in step 904 and token recipient system can receive a DOT from the token transmission system in step 908. Before processing, the token recipient system can validate the public-key digital signatures on the DOT in step 912, to ensure the authenticity of the DOT and the authenticity of the sender. Once validated, the token recipient system can open the cryptographic envelope of the token in step 912 (if it is encrypted) with the private-key that corresponds to the public-key used in securing the token for transmission. This ensures that the token can only be used by the intended token recipient system.

Once validated and decrypted, the token recipient system checks both the token-embedded and external file workflow logic for conflicts in steps 916 and the substeps in 920. If non-critical conflicts are found, resolution may require priority-based assessment or possibly user intervention. The token recipient system can then process the token according to the workflow logic in step 924. As shown in the substeps in 928, processing can include presentation to the user, storage for archival or for later use. At some point, the user of the token recipient system or the system itself can use the received (and possibly stored) token to request one or more operations on the data object to which it refers, as shown in step 932. This processing can be done through the token processing system as shown in step 936.

While the invention has been described in detail, including references to specific embodiments, it will be apparent to one skilled in the art that changes and modifications can be made to the invention without departing from the spirit and scope thereof. Thus, it is intended that the present invention cover the modifications and variations of this invention provided they come within the scope of the appended claims and their equivalents.

What is claimed is:

1. A method of managing one or more data objects within a data processing system using one or more data object tokens, the method comprising the steps of:

generating one or more data objects within a data object origination system;

storing said data objects in a data object storage system;

analyzing said data objects to determine information to be placed in one or more data object tokens;

generating said one or more data object tokens;

transmitting said one or more data object tokens to one or more token recipient systems;

processing said one or more data object tokens; and

processing said one or more data objects based upon said data object tokens.

2. A method as in claim 1, wherein said storing step further comprises storing said data objects in a local storage system.

3. A method as in claim 1, wherein said storing step further comprises storing said data objects in a remote storage system over a network.

4. A method for generating one or more data object tokens associated with a data object, the method comprising:

receiving a token generation request;

validating said token generation request;

analyzing said data object to determine information to be placed in said one or more data object tokens;

populating said data object token with metadata;

creating validation information related to said data object token; and

returning a completed data object token.

5. A method as in claim 4, wherein said validating step further comprises the step of verifying a digital signature using a public key.

6. A method as in claim 4, wherein said analyzing step further comprises the step of determining information to be redacted in said data object.

7. A method as in claim 4, wherein said analyzing step further comprises the step of determining information to be altered in said data object.

8. A method as in claim 4, wherein said populating step further comprises the step of appending one or more optional token extensions.

9. A method for processing one or more data object tokens associated with a data object, the method comprising:

receiving a token processing request;

validating said token processing request;

interpreting said token processing request to determine the operations to be performed on one or more data objects;

commanding said operations to be performed on said one or more data objects by a data object processing system;

validating the results of said operations; and

returning the resulting data object.

10. A method for limiting access by a user to data objects in a data processing system, the method comprising:

storing said data object securely within said data processing system;

creating a data object token corresponding to said data object, said data object containing authentication requirements;

distributing said data object token;

receiving an access request from a data object token user, said request containing data object token user authentication information and access authentication information;

authenticating said access request;

processing the data object to secure it for use by the data object token user; and

transmitting said processed data object to said data object token user.

11. A method as in claim 10, wherein said storing step further comprises performing local encryption on said data object.

12. A method as in claim 10, wherein said storing step further comprises locating said data object behind a firewall.

13. A method as in claim 10, wherein said first validation step further comprises verification of a first public key digital signature, and wherein said second validation step further comprises verification of a second public key digital signature.

14. A method for validating the use of a data object using a data object token in a token based data processing system, the method comprising:

performing a first validation on data object originator authentication information contained within said data object token; and

performing a second validation on data object token user authentication information contained within said data object token.

15. A method as in claim 14, wherein said first validation step further comprises verification of a first public key digital signature, and wherein said second validation step further comprises verification of a second public key digital signature.

16. A method of retrieving a data object in a token-based data processing system, the method comprising the steps of:

receiving a data object token from a token generating system;

storing said data object token;

appending data object token user authentication information;

submitting said data object token to a token processing system; and

receiving said data object.

17. A method as in claim 16, wherein said data object token user information comprises a public key digital signature.

18. A method for distributing analysis and processing activities related to one or more data objects stored in a data object storage system within a token based data processing system, comprising the steps of:

receiving one or more data object tokens in a token recipient system;

storing said data object tokens; and

processing said data object tokens within said token recipient system.

19. A method as in claim 18, wherein said processing step further comprises the steps of:

extracting metadata associated with said one or more data objects from said one or more data object tokens;

performing an analysis on said metadata; and

accessing one or more data objects within said data object storage system based upon said analysis.

20. A method as in claim 18, wherein:

said processing step further comprises a search through data object tokens containing electronic mail header information based on a set of search criteria; and

said accessing step further comprises accessing e-mail messages meeting said search criteria.

21. A token based data processing system that manages one or more data objects within a data processing system using one or more data object tokens, comprising:

a data object origination system for generating one or more data objects;

a data object storage system for storing said one or more data objects;

a token generating system for generating one or more data object tokens;

a token transmission system for transmitting said one or more data object tokens to one or more token recipient systems;

a token processing system for processing said one or more data object tokens; and

a data object processing system for processing said one or more data objects based upon the processing of said one or more data object tokens.

22. A token based data processing system as in claim 21, wherein said data object storage system is contained within said data object origination system.

23. A token based data processing system as in claim 21, wherein said data object storage system is located over a network remotely from said data object origination system.

24. A token based data processing system as in claim 21, wherein said data objects further comprise scanned or imaged documents.

25. A token generating system for generating one or more data object tokens that refer to one or more data objects, comprising:

a means for receiving a token generation request;

a means for validating said token generation request;

a means for analyzing said data object to determine information to be placed in said one or more data object tokens;

a means for populating said data object token with metadata;

a means for creating validation information related to said data object token; and

a means for returning a completed data object token.

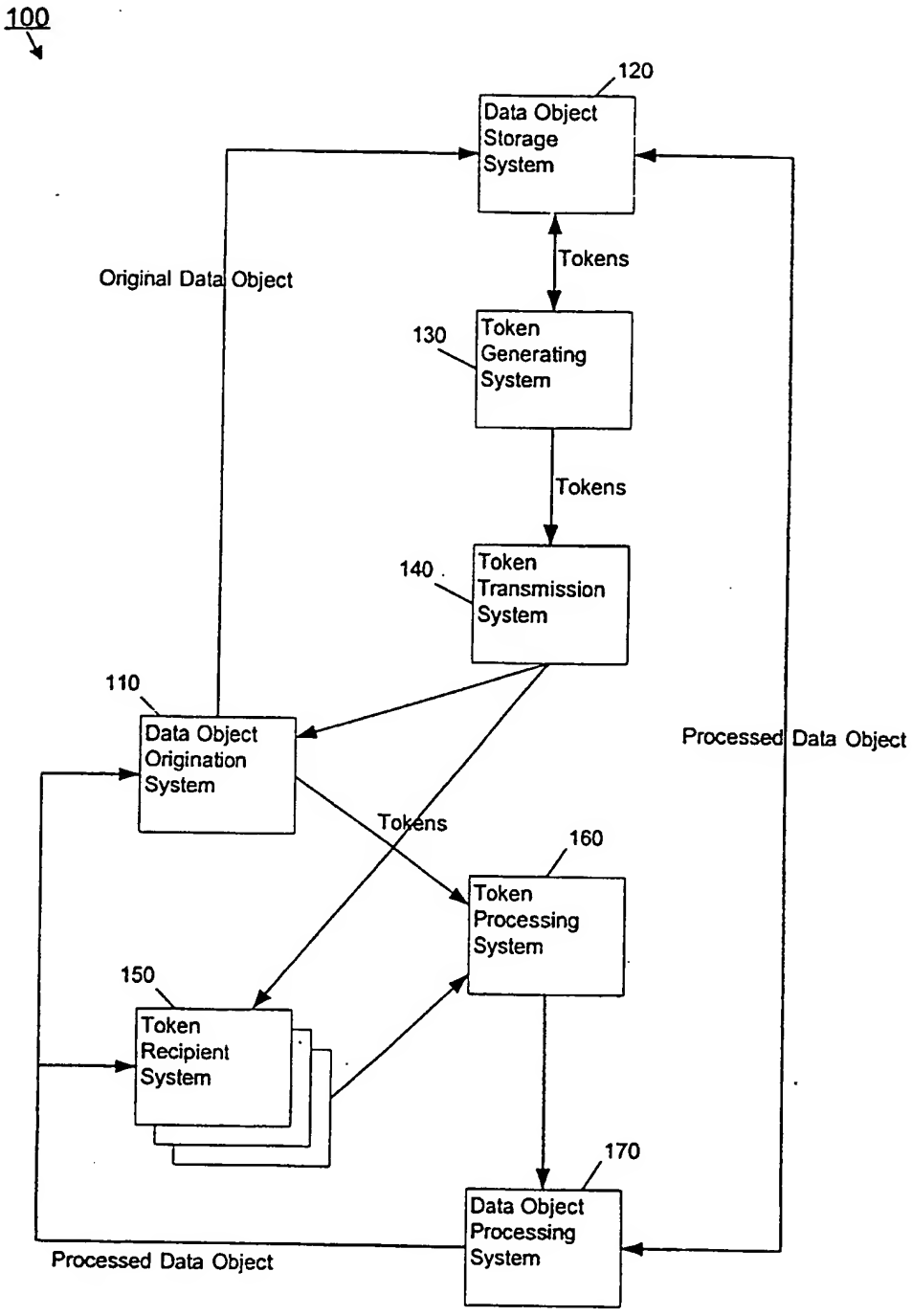

100
120
Data Object Storage System
Tokens
Original Data Object
130
Token Generating System
Tokens
140
Token Transmission System
110
Data Object Origination System
Processed Data Object
Tokens
160
Token Processing System
150
Token Recipient System
170
Data Object Processing System
Processed Data Object

Figure 1

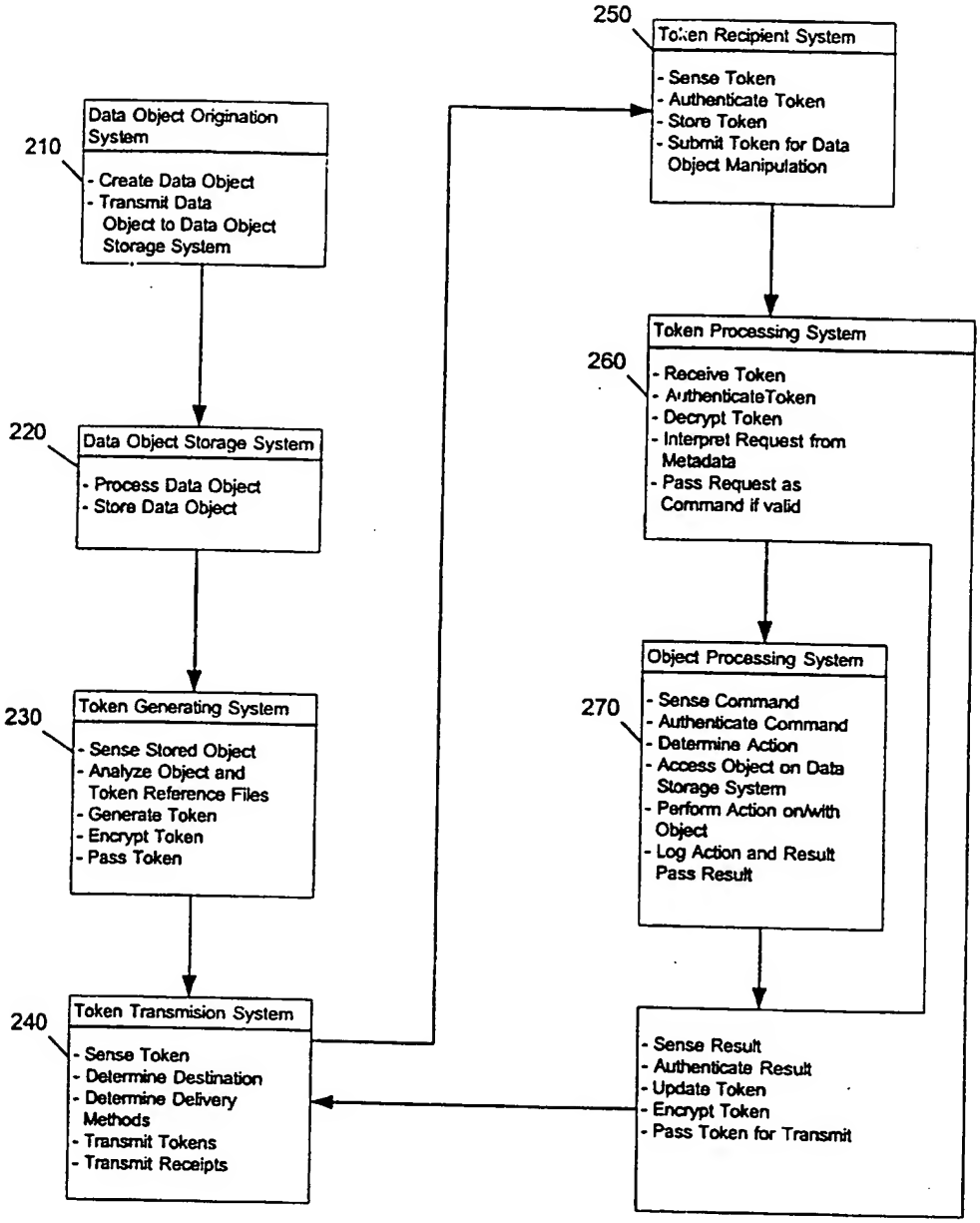

250
Token Recipient System
- Sense Token
- Authenticate Token
- Store Token
- Submit Token for Data Object Manipulation
Data Object Origination System
210
- Create Data Object
- Transmit Data Object to Data Object Storage System
Token Processing System
260
- Receive Token
- AuthenticateToken
- Decrypt Token
- Interpret Request from Metadata
- Pass Request as Command if valid
220
Data Object Storage System
- Process Data Object
- Store Data Object
Object Processing System
270
- Sense Command
- Authenticate Command
- Determine Action
- Access Object on Data Storage System
- Perform Action on/with Object
- Log Action and Result
Pass Result
230
Token Generating System
- Sense Stored Object
- Analyze Object and Token Reference Files
- Generate Token
- Encrypt Token
- Pass Token
240
Token Transmision System
- Sense Token
- Determine Destination
- Determine Delivery Methods
- Transmit Tokens
- Transmit Receipts
- Sense Result
- Authenticate Result
- Update Token
- Encrypt Token
- Pass Token for Transmit

Figure 2

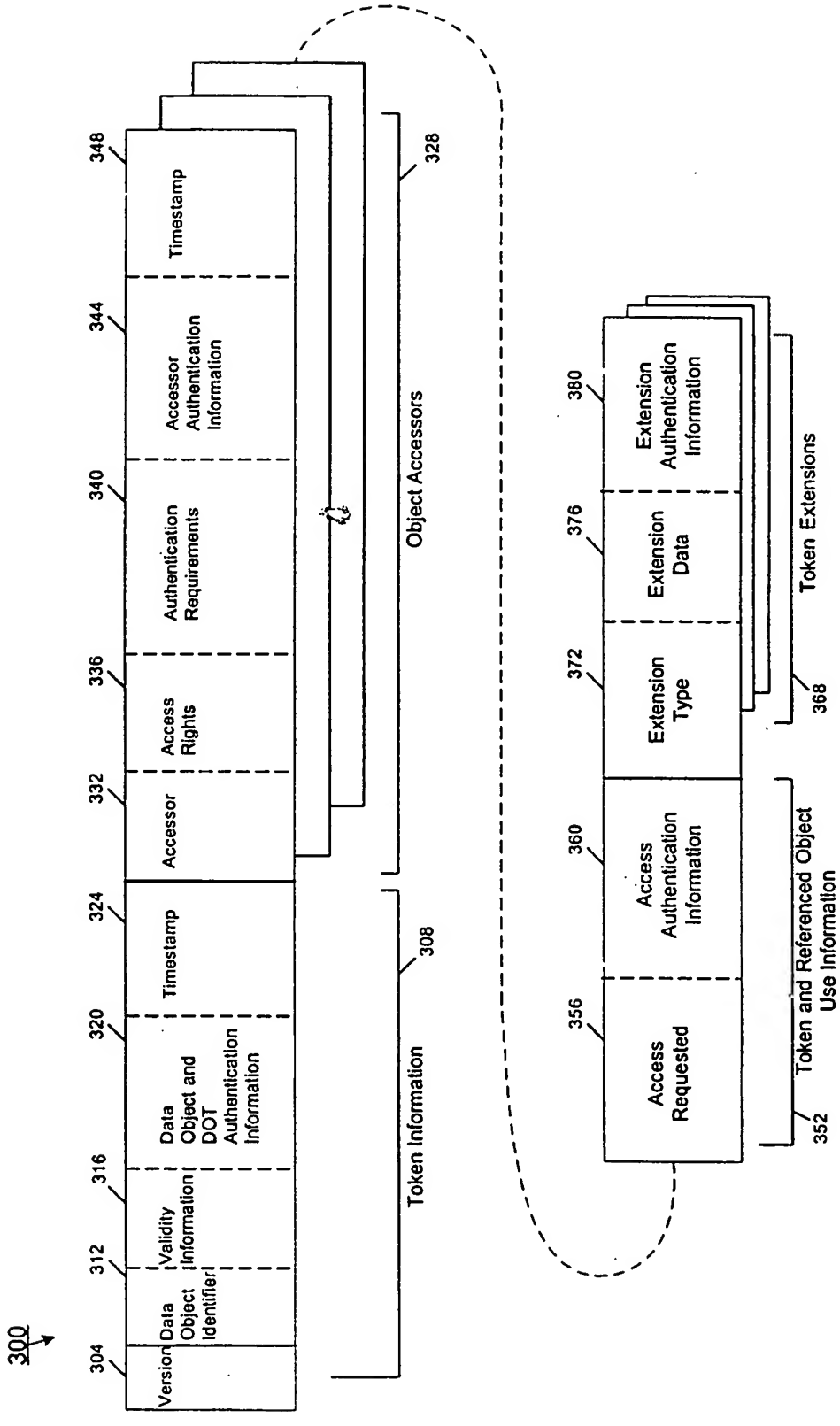

300
304
Version
312
Data Object Identifier
316
Validity Information
320
Data Object and DOT Authentication Information
324
Timestamp
332
Accessor
336
Access Rights
340
Authentication Requirements
344
Accessor Authentication Information
348
Timestamp
328
Object Accessors
308
Token Information
352
Token and Referenced Object Use Information
356
Access Requested
360
Access Authentication Information
372
Extension Type
376
Extension Data
380
Extension Authentication Information
368
Token Extensions

Figure 3

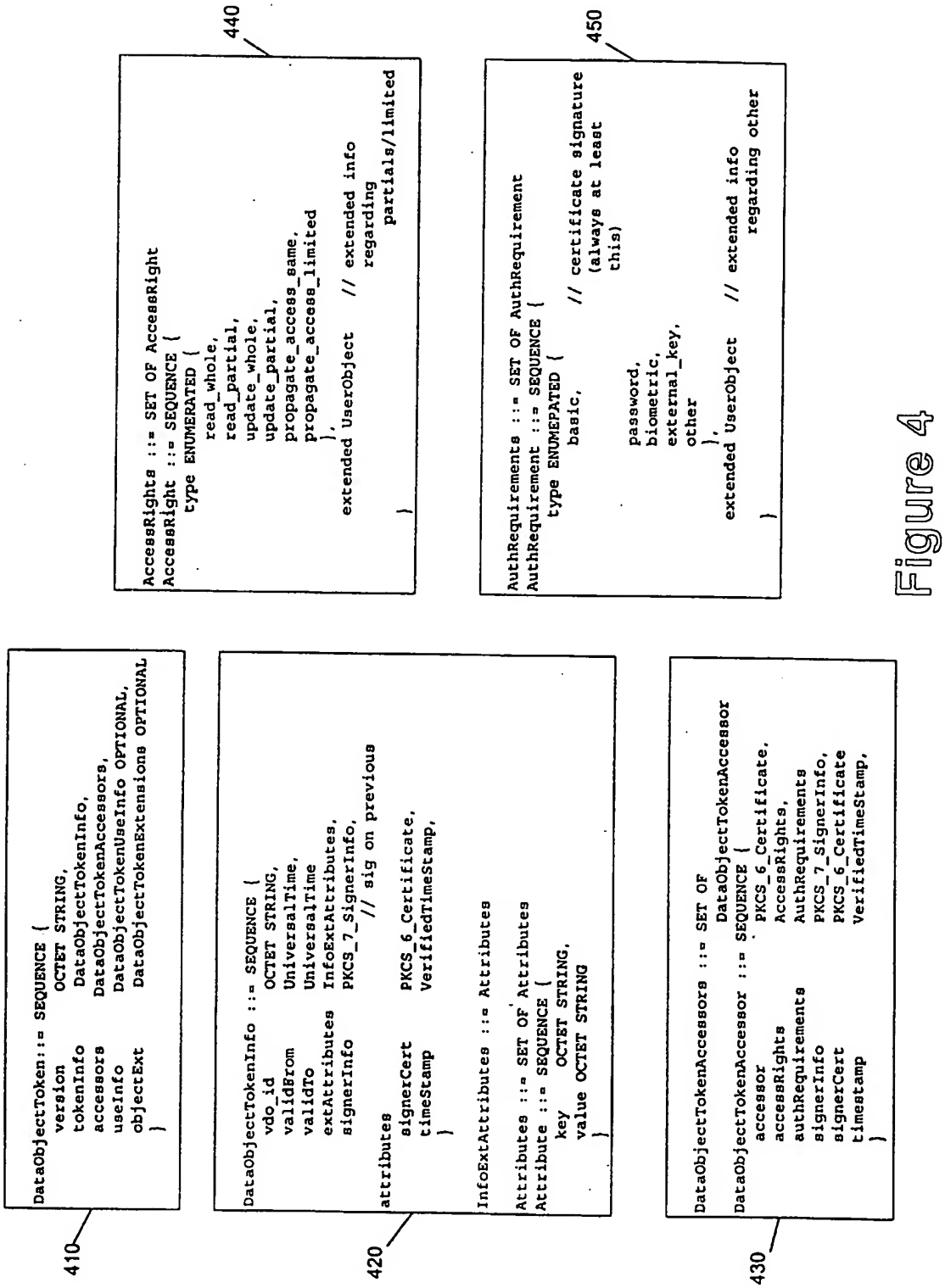

410
DataObjectToken::= SEQUENCE {
version OCTET STRING,
tokenInfo DataObjectTokenInfo,
accessors DataObjectTokenAccessors,
useInfo DataObjectTokenUseInfo OPTIONAL,
objectExt DataObjectTokenExtensions OPTIONAL
}
420
DataObjectTokenInfo ::= SEQUENCE {
vdo_id OCTET STRING,
validFrom UniversalTime,
validTo UniversalTime
extAttributes InfoExtAttributes,
signerInfo PKCS_7_SignerInfo,
// sig on previous
attributes
signerCert PKCS_6_Certificate,
timeStamp VerifiedTimeStamp,
}
InfoExtAttributes ::= Attributes
Attributes ::= SET OF Attributes
Attribute ::= SEQUENCE {
key OCTET STRING,
value OCTET STRING
}
430
DataObjectTokenAccessors ::= SET OF
DataObjectTokenAccessor
DataObjectTokenAccessor ::= SEQUENCE {
accessor PKCS_6_Certificate,
accessRights AccessRights,
authRequirements AuthRequirements
signerInfo PKCS_7_SignerInfo,
signerCert PKCS_6_Certificate
timestamp VerifiedTimeStamp,
}
440
AccessRights ::= SET OF AccessRight
AccessRight ::= SEQUENCE {
type ENUMERATED {
read_whole,
read_partial,
update_whole,
update_partial,
propagate_access_same,
propagate_access_limited
},
extended UserObject // extended info
regarding
partials/limited
}
450
AuthRequirements ::= SET OF AuthRequirement
AuthRequirement ::= SEQUENCE {
type ENUMERATED {
basic, // certificate signature
(always at least
this)
password,
biometric,
external_key,
other
},
extended UserObject // extended info
regarding other
}

Figure 4

510

```
DataObjectTokenUseInfo ::= SEQUENCE {
      accessRequested       Access,
      authentication        UserObject,
      accessorInfo                PKCS_7_SignerInfo,
      accessorCert                PKCS_6_Certificate
      }

Access ::= SEQUENCE {
      type ENUMERATED {
            read_whole,
            read_partial,
            update_whole,
            update_partial
            },
      extended UserObject
            }
```

520

```
DataObjectTokenExtensions ::= SET OF DataObjectTokenExtension
DataObjectTokenExtension ::= SEQUENCE {
      extType           OCTET STRING,
      extData           UserObject,
      signerInfos       PKCS_7_SignerInfos,
      signerCerts       PKCS_6_Certificates
      }

PKCS_7_SignerInfos ::= SET OF PKCS_7_SignerInfo
PKCS_6_Certificates ::= SET OF PKCS_6_Certificate
```

Figure 5

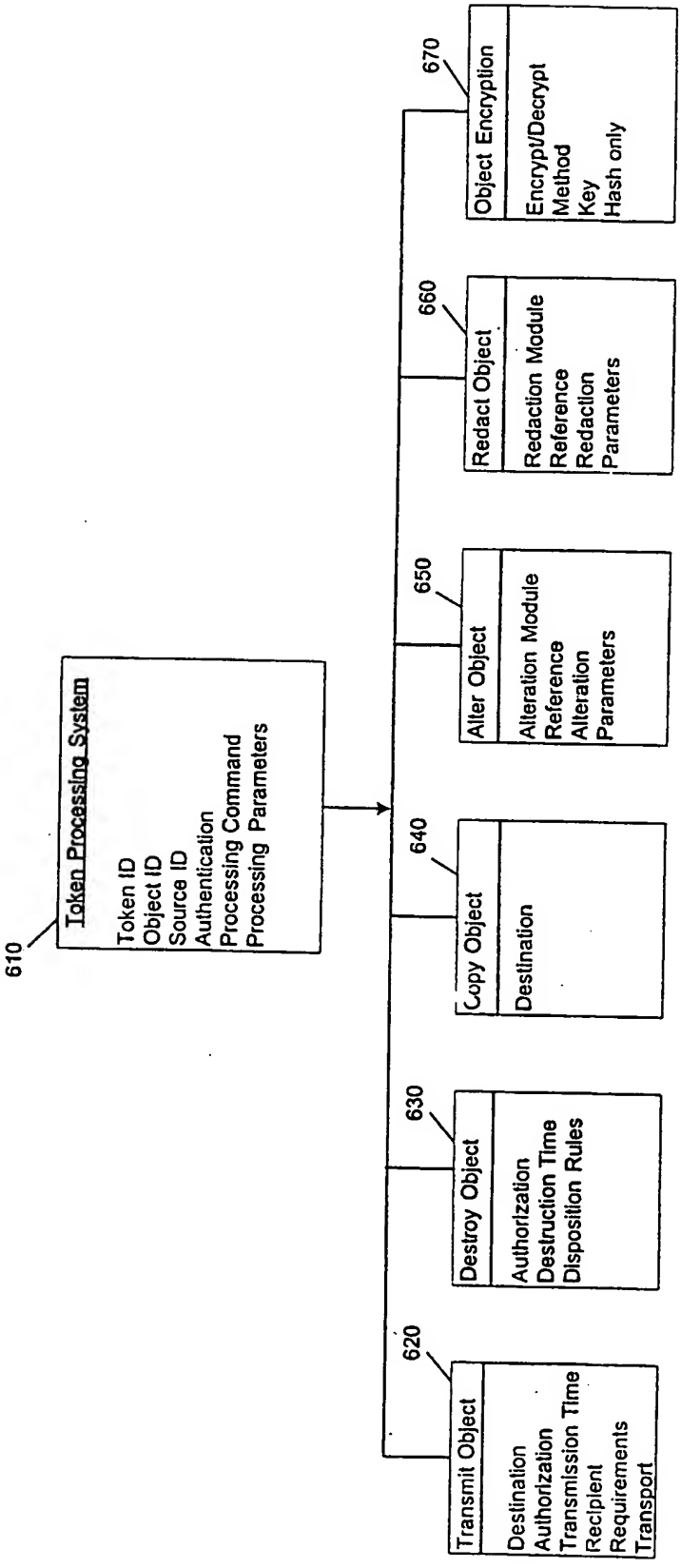


Figure 6

Partial Token Based Data Processing System Command Table

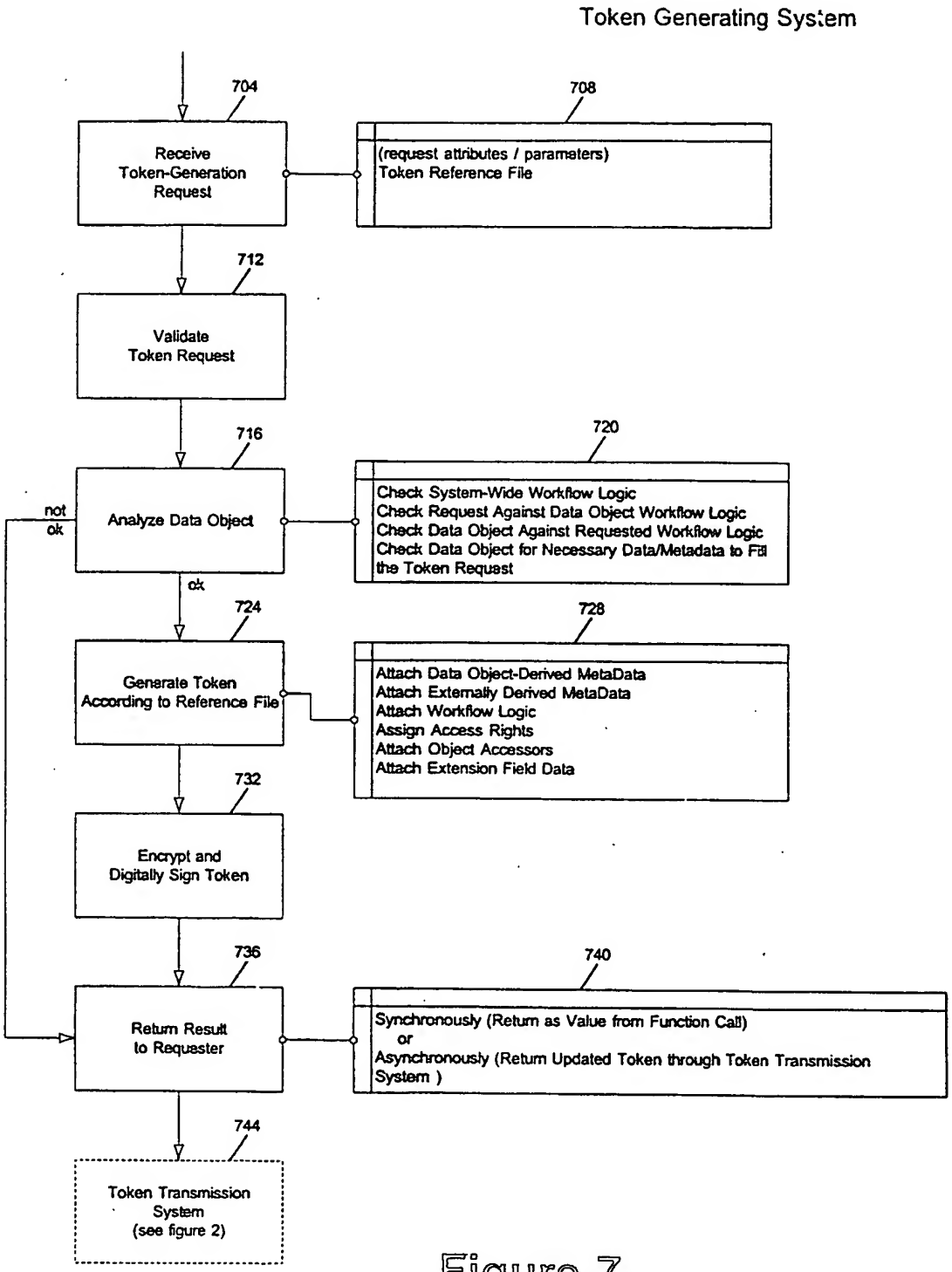

Token Generating System
704
Receive Token-Generation Request
708
(request attributes / parameters)
Token Reference File
712
Validate Token Request
716
Analyze Data Object
not ok
ok
720
Check System-Wide Workflow Logic
Check Request Against Data Object Workflow Logic
Check Data Object Against Requested Workflow Logic
Check Data Object for Necessary Data/Metadata to Fil the Token Request
724
Generate Token According to Reference File
728
Attach Data Object-Derived MetaData
Attach Externally Derived MetaData
Attach Workflow Logic
Assign Access Rights
Attach Object Accessors
Attach Extension Field Data
732
Encrypt and Digitally Sign Token
736
Return Result to Requester
740
Synchronously (Return as Value from Function Call)
or
Asynchronously (Return Updated Token through Token Transmission System)
744
Token Transmission System
(see figure 2)

Figure 7

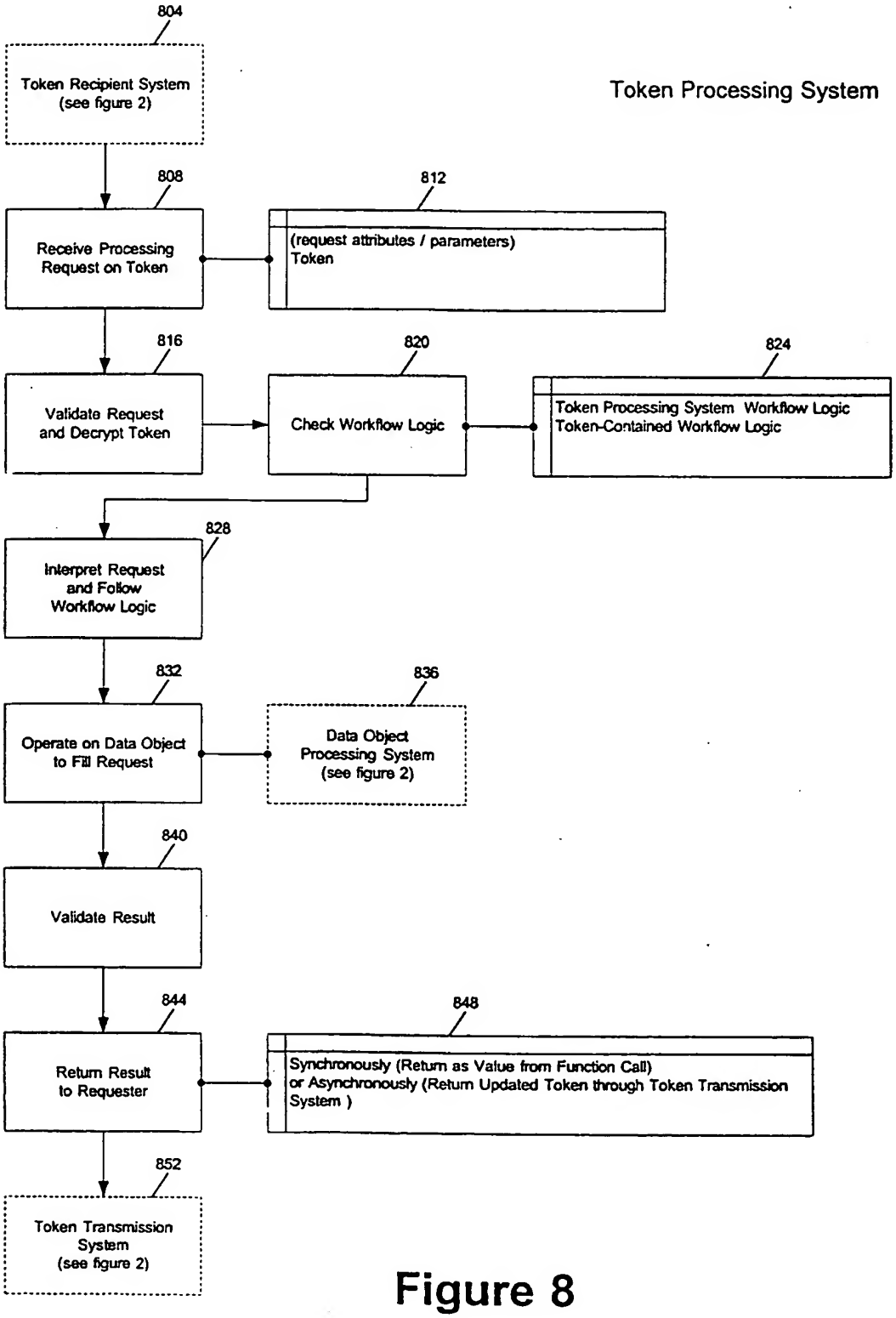

804
Token Recipient System
(see figure 2)
Token Processing System
808
Receive Processing
Request on Token
812
(request attributes / parameters)
Token
816
Validate Request
and Decrypt Token
820
Check Workflow Logic
824
Token Processing System Workflow Logic
Token-Contained Workflow Logic
828
Interpret Request
and Follow
Workflow Logic
832
Operate on Data Object
to Fill Request
836
Data Object
Processing System
(see figure 2)
840
Validate Result
844
Return Result
to Requester
848
Synchronously (Return as Value from Function Call)
or Asynchronously (Return Updated Token through Token Transmission
System)
852
Token Transmission
System
(see figure 2)

Figure 8

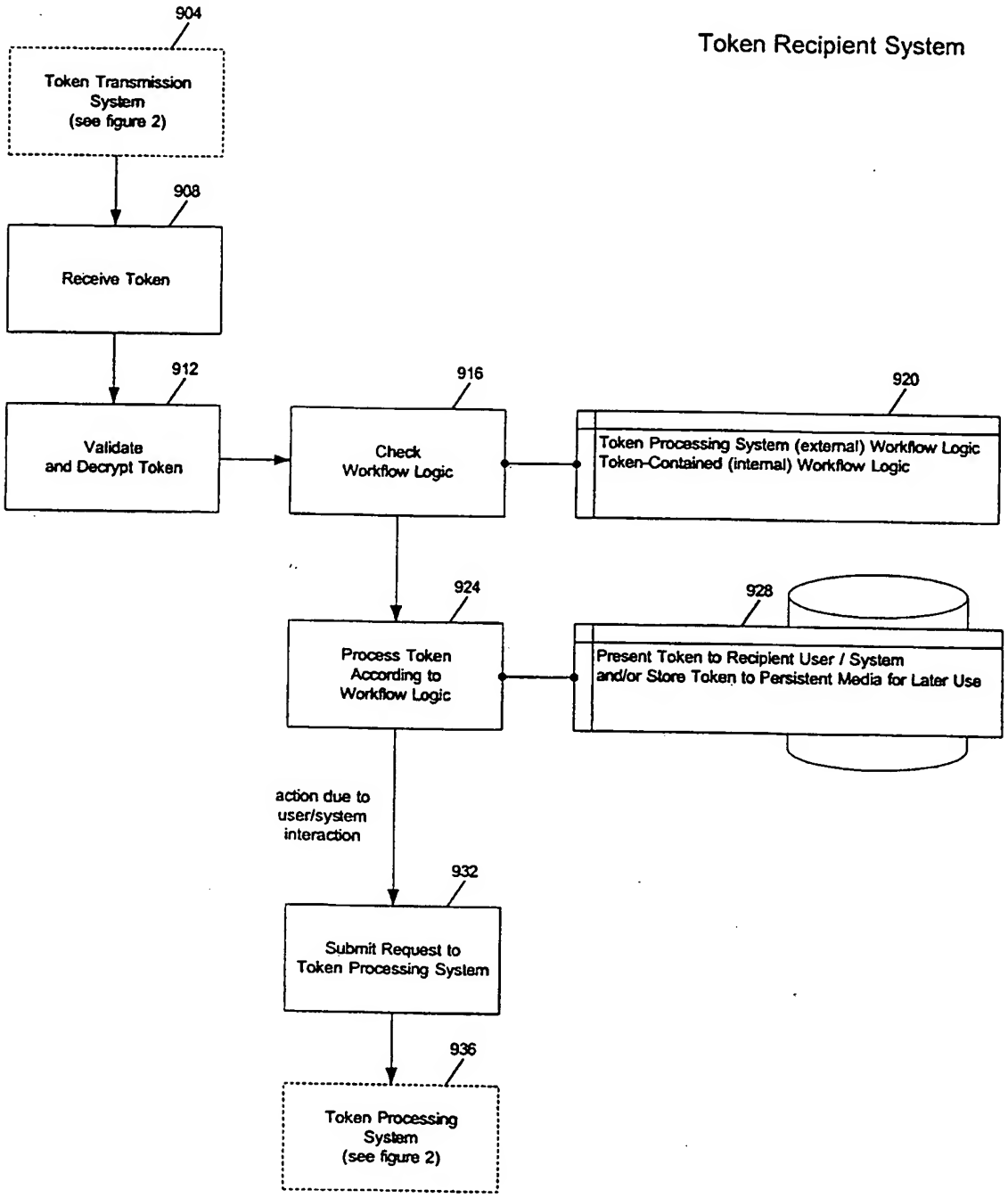

Token Recipient System
904
Token Transmission System (see figure 2)
908
Receive Token
912
Validate and Decrypt Token
916
Check Workflow Logic
920
Token Processing System (external) Workflow Logic
Token-Contained (internal) Workflow Logic
924
Process Token According to Workflow Logic
928
Present Token to Recipient User / System and/or Store Token to Persistent Media for Later Use
action due to user/system interaction
932
Submit Request to Token Processing System
936
Token Processing System (see figure 2)

Figure 9